Rainfall data analysis using wavelet transform

CELSO AUGUSTO GUIMARÃES SANTOS
Department of Civil Construction Technology, Federal University of Paraíba, 58059-900 João Pessoa, Paraíba, Brazil
celso@dec.ufpb.br

CARLOS OLIVEIRA GALVÃO
Department of Civil Engineering, Federal University of Campina Grande, Av. Aprígio Veloso 882, 58109-970 Campina Grande, Paraíba, Brazil

RICARDO MACHADO TRIGO
Department of Physics, University of Lisbon, Campo Grande, Ed C8, Piso 6, P-1749-016 Lisbon, Portugal

Abstract An application of wavelet analysis is done with a long time series of the total monthly rainfall from several places from different regions of the world (i.e. Iberian Peninsula, Japan and north-eastern Brazil). Such an analysis was performed in order to fully characterize the distinct time–frequency rainfall variability observed in each of these areas. Besides the rainfall variability analysis, the main frequency components in the time series are studied by the global wavelet spectrum, revealing how the monthly rainfall frequency of each place is composed. This analysis is considered to be more accurate than the standard Fourier analysis. The modulation in separated bands was done in order to extract additional information; e.g. the 8–16-month band was examined by an average of all scales between 8 and 16 months, giving a measure of the average monthly variance versus time, where the periods with low or high variance could be identified.

Key words multiscale analysis; rainfall data; wavelet

INTRODUCTION

The wavelet transform is a recent advance in signal processing that has attracted much attention since its theoretical development in 1984 by Grossman & Morlet (1984). Its use has increased rapidly as an alternative to the Fourier Transform in preserving local, non-periodic, multiscaled phenomena. It has advantages over classical spectral analysis, because it allows the analysis of different scales of temporal variability and it does not need a stationary series. Thus, it is appropriate to analyse irregular distributed events and time series that contain non-stationary power at many different frequencies.

Several applied fields are making use of wavelets such as astronomy, acoustics, data compression, nuclear engineering, sub-band coding, signal and image processing, neurophysiology, music, magnetic resonance imaging, speech discrimination, optics, fractals, radar, human vision, pure mathematics, and geophysics such as tropical convection, the El Niño-Southern Oscillation, atmospheric cold fronts, temperature variability, the dispersion of ocean waves, wave growth and breaking, structures in turbulent flows, and stream flow characterization (Graps, 1995; Torrence & Compo, 1998; Farge, 1992; Smith *et al.*, 1998).

WAVELET TRANSFORM

Mathematical transformations are applied to signals to obtain further information from that signal that is not readily available in the raw signal. There are several transformations that can be applied, among which the Fourier transforms are probably by far the most popular.

Wavelet analysis maintains time and frequency localization in a signal analysis by decomposing or transforming a one-dimensional time series into a diffuse two-dimensional time-frequency image simultaneously. Then, it is possible to get information on both the amplitude of any "periodic" signals within the series, and how this amplitude varies with time.

An example of a basic wave or mother wavelet, as it is known in the literature, is the Morlet wavelet. This "wavelet" has the advantage of incorporating a wave of a certain period, as well as being finite in extent. Assuming that the total width of this wavelet is about 10 years, it is possible to find the correlation between this curve and the first 10 years of the time series as shown later in Fig. 1(a). This single number gives a measure of the projection of this wave packet on the data during the 1890–1900 period, i.e. how much [amplitude] does the 10-year period resemble a sine wave of this width [frequency]. By sliding this wavelet along the time series, a new time series of the projection amplitude *vs* time can be constructed. Finally, the "scale" of the wavelet can be varied by changing its width.

RAINFALL DATA

Monthly rainfall data from five different stations characteristic of three distinct regions of the world were used: Lisbon, Madrid and Barcelona in the Iberian Peninsula, Matsuyama in Japan, and Angicos in north-eastern Brazil. Each location has its own distinct hydroclimatic characteristics, including seasonal distribution of monthly rainfall and temperature, as well as low frequency oscillations (interannual or even interdecadal) and possible trends. All these specific time series characteristics can be analysed with the wavelets tool.

DATA ANALYSIS

Since the present data are distributed monthly, the parameters for the wavelet analysis are set as the time interval $\delta t = 1$ month, the start scale $s_0 = 2$ months, the scale width $\delta j = 0.25$, which will do four sub-octaves per octave, and there will be seven powers-of-two.

Wavelet power spectrum

Figures 1(b), 2(b), 3(b), 4(b) and 5(b) show the power (absolute value squared) of the wavelet transform for the monthly rainfall in Matsuyama, Lisbon, Madrid, Barcelona and Angicos, respectively. The (absolute value)^2 gives information on the relative power at a certain scale and a certain time. These figures show the actual oscillations

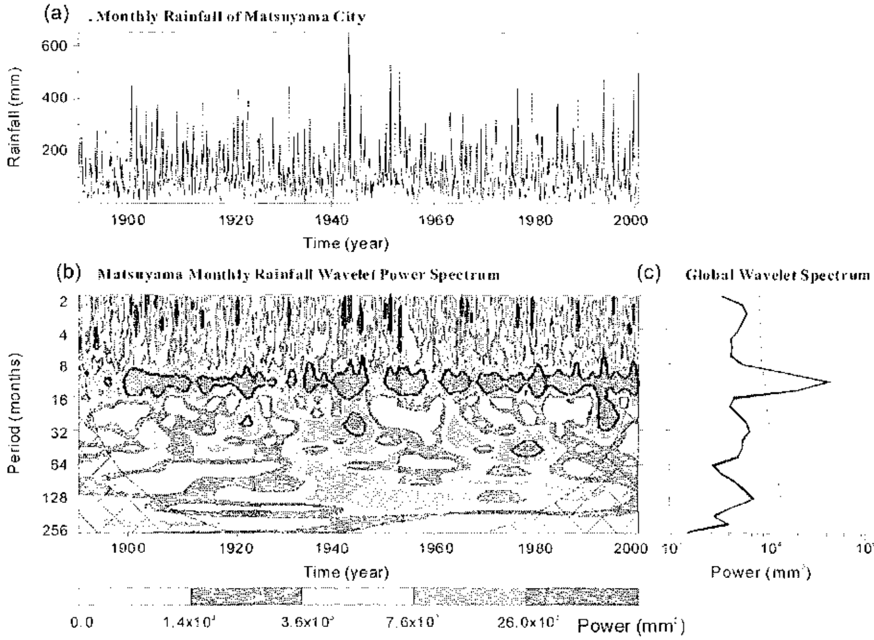

Fig. 1 (a) Monthly rainfall of Matsuyama city. (b) The wavelet power spectrum using Morlet mother wavelet. The contour levels are chosen so that 75%, 50%, 25%, and 5% of the wavelet power is above each level, respectively. The cross-hatched region is the cone of influence, where zero padding has reduced the variance. Black contour is the 5% significance level, using a white-noise background spectrum. (c) The global wavelet power spectrum (black line). The dashed line is the significance for the global wavelet spectrum, assuming the same significance level and background spectrum as in (b).

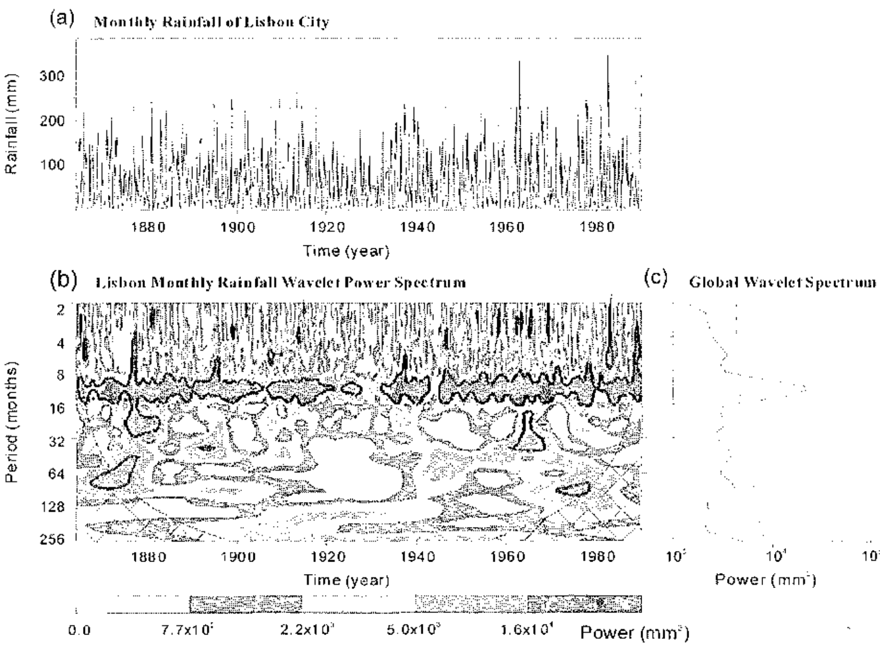

Fig. 2 As in Fig. 1 but for Lisbon precipitation.

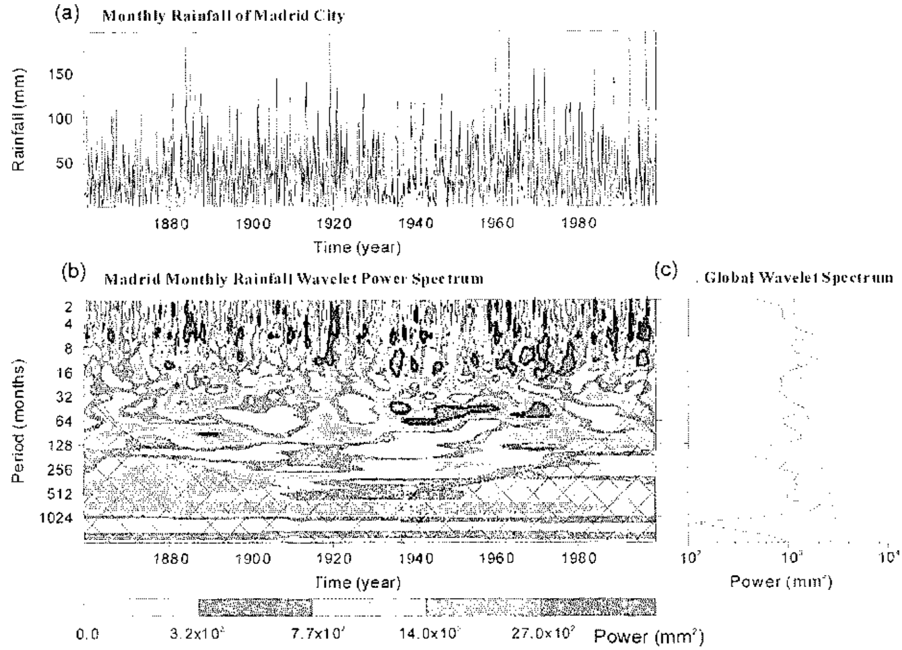

Fig. 3 As in Fig. 1 but for Madrid precipitation.

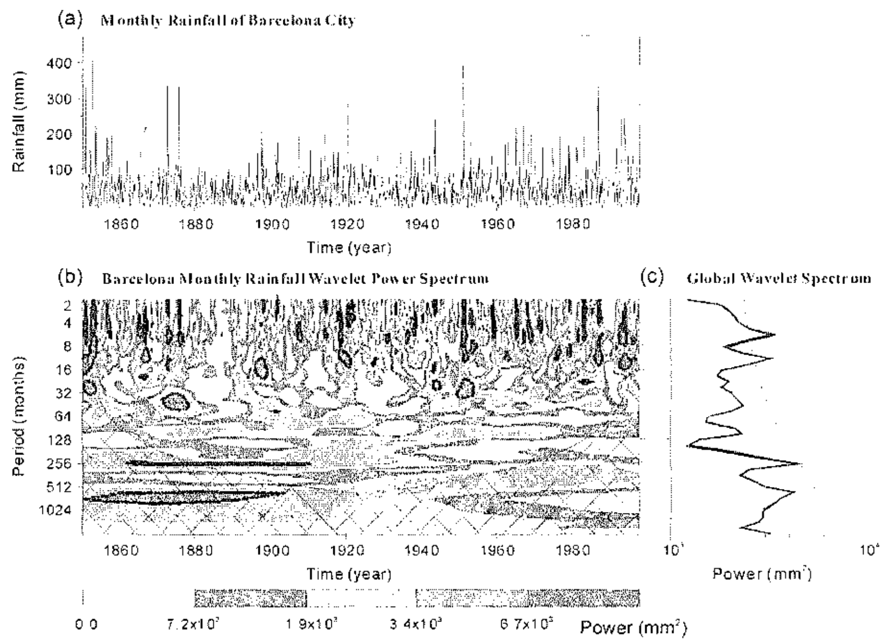

Fig. 4 As in Fig. 1 but for Barcelona precipitation.

of the individual wavelets, rather than just their magnitude. Observing these figures, the concentration of power can be easily identified in the frequency or time domain.

The cross-hatched regions in these figures are the cone of influence, where zero padding has reduced the variance. Because we are dealing with finite-length time series, errors will occur at the beginning and end of the wavelet power spectrum (Santos *et al.*, 2001).

The black contours in the same figures are the 5% significance level, using a white-noise background spectrum. The exception is Angicos (Fig. 5) where we have used a red-noise background spectrum instead of white-noise.

The null hypothesis is defined for the wavelet power spectrum as assuming that the time series has a mean power spectrum; if a peak in the wavelet power spectrum is significantly above this background spectrum, then it can be assumed to be a true feature with a certain percent confidence. For definitions, "significant at the 5% level" is equivalent to "the 95% confidence level," and implies a test against a certain background level, while the "95% confidence interval" refers to the range of confidence about a given value. The 95% confidence implies that 5% of the wavelet power should be above this level. More details can be found in Santos *et al.* (2001).

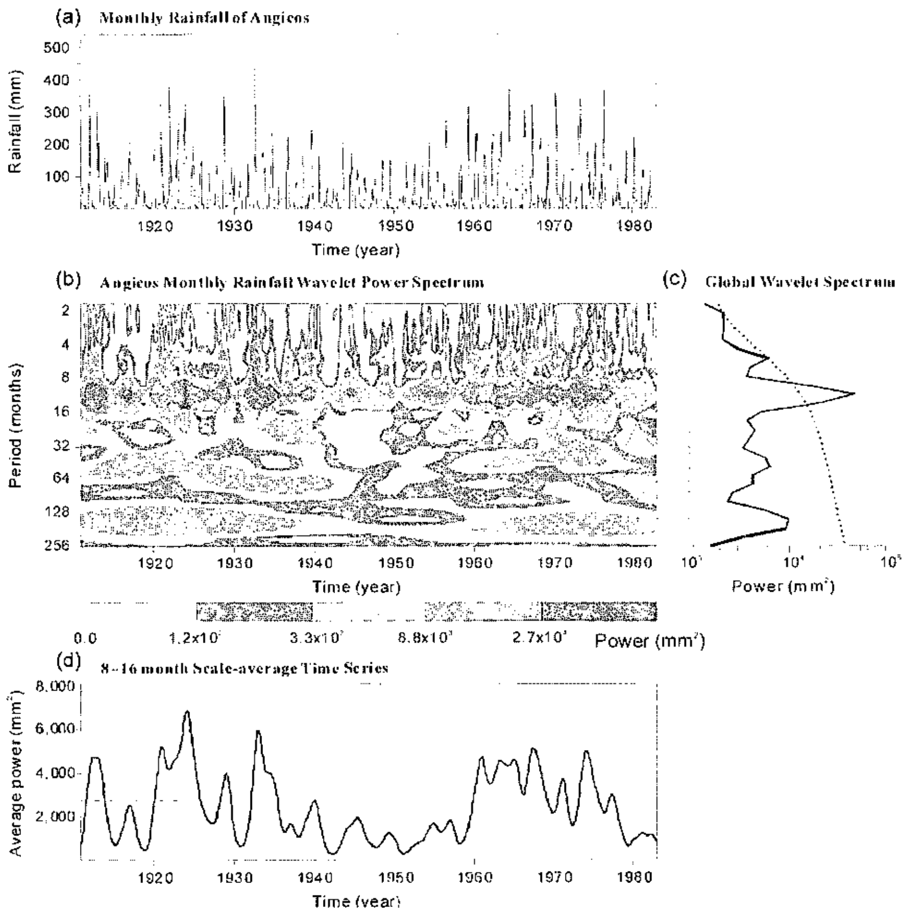

Fig. 5 As in Fig. 1 but for Angicos precipitation, and using a red-noise (autoregressive $\alpha = 0.4891$) background spectrum in (b) and (c).

Global wavelet power spectrum

The annual frequency (periodicity at 12 months) or semi-annual frequency (periodicity at 6 months) of these time series are confirmed by an integration of power over time (Figs 1(c), 2(c), 3(c), 4(c) and 5(c)), which show the significant peaks above the 95% confidence level for the global wavelet spectrum, assuming white-noise (Figs 1(c), 2(c), 3(c) and 4(c)) or red noise (with $\alpha = 0.4891$, Fig. 5(c)), represented by the dashed lines. These global wavelet spectra provide an unbiased and consistent estimation of the true power spectrum of the time series, and thus they are a simple and robust way to characterize the time series variability. Global wavelet spectra should be used to describe rainfall variability in non-stationary hyetographs. For regions that do not display long-term changes in hyetograph structures, global wavelet spectra are useful for summarizing a region's temporal variability and comparing it with rainfall in other regions.

Scale-average time series

The scale-average wavelet power is a time series of the average variance in a certain band. In the case of Fig. 5(d), it is the 8- to 16-months band. It is used to examine modulation of one time series by another, or modulation of one frequency by another within the same time series. This figure is made by the average of Fig. 5(b) over all scales between 8 and 16 months, which gives a measure of the average year variance versus time. Because Angicos presents a strong concentration of power between the 8- to 16-month band, important reductions of power in this band correspond to dry periods such as the 1930–1960 case. The variance plot shows a distinct period when monthly rainfall variance was low, i.e. from 1930 to 1960.

CONCLUSIONS

To study the variability of the monthly rainfall time series in the Iberian Peninsula, Japan and north-eastern Brazil, wavelet analysis was applied. The wavelet power spectra show big power concentrations between the 8- to 16-month bands for all the regions studied, revealing an annual periodicity of such events, which is confirmed by the peaks of the integration of transform magnitude vectors over time that again show a strong annual signal. For Madrid and Barcelona, they also showed a semi-annual periodicity. Non-significant low frequency oscillations, related to the North Atlantic Oscillation, were identified for the Iberian stations. Periods with low variance in the 8- to 16-month band were identified for Angicos, coincident with one of the major droughty events in that semiarid part of Brazil. Wavelet transform revealed important features of the rainfall time series. These results encourage its use, instead of the original series, in hydrological modelling studies.

Acknowledgements The writers wish to thank Dr Christopher Torrence of the Advanced Study Program at National Center for Atmospheric Research, Colorado, USA for providing the wavelet analysis computer program.

REFERENCES

Farge, M. (1992) Wavelet transforms and their applications to turbulence. *Ann. Rev. Fluid Mech.* **24**, 395–457.

Graps, A. (1995) An introduction to wavelets. *IEEE Comp. Sci. Engng* **2**(2), 50–61.

Grossman, A. & Morlet, J. (1984) Decomposition of Hardy functions into square integrable wavelets of constant shape. *SIAM J. Math. Anal.* **15**, 723–736.

Santos, C. A. G., Galvão, C. O., Suzuki, K. & Trigo, R. M. (2001) Matsuyama city rainfall data analysis using wavelet transform. *Ann. J. Hydraul. Engng JSCE*, **45**, 211–216.

Smith, L.C., Turcitte, D. L. & Isacks, B.L. (1998) Stream flow characterization and feature detection using a discrete wavelet transform. *Hydrol. Processes* **12**, 233–249.

Torrence, C. & Compo, G. P. (1998) A practical guide to wavelet analysis. *Bull. Am. Met. Soc.* **79**(1), 61–78.